PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 13/26	A1	(11) International Publication Number: WO 99/30243 (43) International Publication Date: 17 June 1999 (17.06.99)

(21) International Application Number: PCT/US98/23354

(22) International Filing Date: 2 November 1998 (02.11.98)

(30) Priority Data:
08/988,233 10 December 1997 (10.12.97) US

(71) Applicant *(for all designated States except US)*: INTEL CORPORATION [US/US]; 2200 Mission College Boulevard, Santa Clara, CA 95052 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PAWLOWSKI, Stephen, S. [US/US]; 6624 S.W. 158th Avenue, Beaverton, OR 97007 (US). LAU, Daniel, G. [US/US]; 1800 Austin Avenue, Los Altos, CA 94024 (US).

(74) Agents: TAYLOR, Edwin, H. et al.; Blakely, Sokoloff, Taylor & Zafman LLP, 7th floor, 12400 Wilshire Boulevard, Los Angeles, CA 90025 (US).

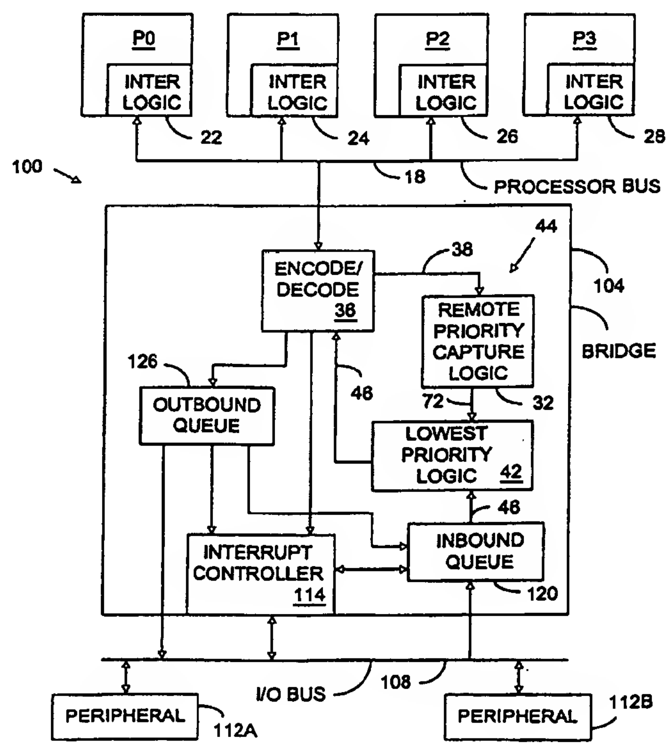
(81) Designated States: AL, AM, AT, AT (Utility model), AU (Petty patent), AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MECHANISM THAT PERFORMS INTERRUPT DESTINATION REDIRECTION

(57) Abstract

One embodiment of the invention includes an apparatus (10), such as a bridge, for use in connection with computer system. The apparatus includes remote priority capture logic (32) to hold task priority data indicative of a task priority of each processor P0 in the computer system that is available for lowest priority interrupt destination arbitration (LPIDA). The apparatus also includes lowest priority logic (42) to perform the LPIDA to select processor in the computer system is to receive an interrupt message based on contents of the remote priority capture logic (32). Another embodiment of the invention includes a multi-processor system having processors and a processor bus (18) coupled to the processors (P0, P1, P2, P3). The system includes remote priority capture logic (32) to hold task priority data indicative of a task priority of the processors while they are available for lowest priority interrupt destination arbitration (LPIDA). The system also includes lowest priority logic (42) to perform the LPIDA to select processors (P0, P1, P2, P3) is to receive an interrupt message based on contents of the remote priority capture logic (32), the interrupt message being provided to the processor through the processor bus (18).



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MECHANISM THAT PERFORMS INTERRUPT DESTINATION REDIRECTION

Related Application: The present application and application no. __/___,___, entitled "Transactions Supporting Interrupt Destination Redirection and Level Triggered Interrupt Semantics" (Docket No. 42390.P4958), which is filed concurrently with the present application, include overlapping disclosures but claim different subject matter.

Background of the Invention

Technical Field of the Invention: The present invention relates to computers and, more particularly, to a mechanism that may perform interrupt destination redirection in a computer system.

Background Art: Processors such as the Pentium® processor and the Pentium® Pro processor manufactured by Intel Corporation are often used in multi-processor systems. Various devices including input and/or output (I/O) devices and other processors may seek to interrupt a processor. To interrupt a processor, an I/O device provides a signal to an interrupt controller, which in turn presents an interrupt request to the processor.

In the case of the Pentium® processor and Pentium® Pro processor, the interrupt controller communicates interrupt information to the processors through a three-wire serial bus, called an APIC (Advanced Programmable

Interrupt Controller) bus. The APIC serial bus includes two data conductors and a clock signal conductor.

The Pentium® processor and Pentium® Pro processor include an internal APIC. The APIC includes a local mask register called a Task Priority Register (TPR) that has 8 bits to designate up to 256 priority states, although some of them are reserved. The contents of the TPR is changed to reflect the level of priority of the tasks being performed by the processor.

A lowest priority interrupt is one that although directed to a particular processor, may be redirected to a processor in a group of processors having the lowest priority in its TPR. The arbitration process involves comparing the 8 bits of the TPR of each processor participating in the arbitration. The bits of each processor are asserted one bit at a time, beginning with the most significant bit (MSB), onto the APIC bus line, which is connected in an open drain arrangement to each of the processors. The bits are inverted onto the APIC bus line so that a low voltage (0) has a higher priority that a high voltage (1). First, the MSB from the TPR of each processor participating in the arbitration is asserted on the APIC bus line. If any of the processors asserts a low voltage on the APIC bus line, the line is pulled low. A processor asserting a high voltage discovers there is another processor with a lower priority if the APIC bus line is pulled low. The processor drops out of consideration if another processor has a lower priority. Then, the second MSB from the TPR of each remaining processor is asserted on the APIC

bus line. If a processor asserts a high voltage as the second MSB, but the line is pulled low, the processor drops out of consideration. The third MSB and later the fourth MSB of each remaining processor are asserted on the APIC bus line in similar fashion and so forth to the least significant bit (LSB). If two or more processors have equal priorities after all eight bits have been asserted, the processor with the lowest local APIC identification (ID) number is chosen to receive the interrupt vector. The local APIC ID number is assigned at power up.

There are certain disadvantages with the APIC serial bus. First, the serial bus is poor at voltage scaling between the interrupt controller (e.g., 3.3 volts) and the processor (e.g., 2.5 or 1.8 volts). It is difficult for provide transistors in a processor that interface between such disparate voltages. As the voltage of the processor core decreases with new generations of processors, the problem will be even greater.

Second, the frequency of the processor core (e.g., often much greater than 200 MHz) is much greater than the frequency of the APIC serial bus (e.g., 16 MHz). As processor frequencies increase, the problem will be even greater. It is difficult to interface between such disparate frequencies. The problem is greater because the signals are independent of each other.

Third, the APIC serial bus is relatively slow. In some implementations, it takes roughly 2 to 3 microseconds to deliver an interrupt. As more I/O intensive

functions are used, the speed at which the serial bus can deliver interrupts becomes limiting.

The present invention is directed to over coming or reducing the effect of one or more of the above-recited problems with the APIC serial bus.

Summary of the Invention

One embodiment of the invention includes an apparatus for use in connection with computer system. The apparatus includes remote priority capture logic to hold task priority data indicative of a task priority of each processor in the computer system that is available for lowest priority interrupt destination arbitration (LPIDA). The apparatus also includes lowest priority logic to perform the LPIDA to select which processor in the computer system is to receive an interrupt message based on contents of the remote priority capture logic.

Another embodiment of the invention includes a multi-processor system having processors and a processor bus coupled to the processors. The system includes remote priority capture logic to hold task priority data indicative of a task priority of the processors while they are available for lowest priority interrupt destination arbitration (LPIDA). The system also includes lowest priority logic to perform the LPIDA to select which of the processors is to receive an interrupt message based on contents of the remote priority capture

logic, the interrupt message being provided to the processor through the processor bus.

Brief Description of the Drawings

The invention will be understood more fully from the detailed description given below and from the accompanying drawings of embodiments of the invention which, however, should not be taken to limit the invention to the specific embodiments described, but are for explanation and understanding only.

FIG. 1 is a block diagram representation of a multi-processor system including lowest priority logic for directing interrupts to a lowest priority processor.

FIG. 2 is a block diagram representation of an example of certain details of one embodiment of the processors of the system of FIG 1.

FIG. 3 is a block diagram representation of an example of certain details of one embodiment of the remote priority capture logic and lowest priority logic of FIG. 1.

FIG. 4 is an illustration of one embodiment of a remote task priority register (RTPR) in the remote priority capture logic of FIG. 3.

FIG. 5 is a block diagram representation of one embodiment of a multi-processor system including interrupt direction logic, remote priority capture logic,

and encode/decode logic in a bridge for directing interrupts to a lowest priority processor.

FIG. 6 is a block diagram representation of a multi-processor system similar to that of FIG. 5 with the addition of an APIC serial bus.

FIG. 7 illustrates a two phase special cycle for RTPR update.

Detailed Description of Preferred Embodiments

Referring to FIG. 1, a multi-processor computer system 10 includes processors P0, P1, P2, and P3 connected through a processor bus 18. In some embodiments, processor bus 18 is referred to as a front side bus. The invention may be used in connection with a system having more or less than four processors. Processors P0, P1, P2, and P3 include interrupt control logic 22, 24, 26, and 28, respectively, which include a task priority designation that indicates a task priority, which is the priority level at which interrupts are serviced. As an example, the task priority designation may be an 8-bit number. Typically, if the priority of the interrupt is lower than the value in the task priority register of the processor, then the processor will not act on the interrupt.

Remote priority capture logic 32 holds task priority data that are indicative of task priorities of those of processors P0, P1, P2, and P3 that are available for lowest priority interrupt destination arbitration (LPIDA). For example, the task priority data may be the 4 MSBs of the task priority

designation of one or more of processors P0, P1, P2, and P3. As used herein, the term "remote" means off a processor die. Under one embodiment, remote priority capture logic 32 also holds task priority data that is indicative of task priorities of those of processors P0, P1, P2, and P3 that are not available for LPIDA, but are operating in system 10. The task priority data may be assembled in remote priority capture logic 32 as follows. Signals representative of the task priority of one or more of processors P0, P1, P2, and P3 are provided by the processor(s) to processor bus 18. Encode/decode logic 36 decodes these signals from processor bus 18 and provides signals responsive to them to remote priority capture logic 32 through conductors 38. There is not necessarily a one-to-one correspondence between bits of the signals on processor bus 18, conductors 38, and remote priority capture logic 32. For example, the bits could be inverted.

A write cycle signal including a lowest priority interrupt message is received by lowest priority logic 42 through conductors 46. Lowest priority logic 42 performs LPIDA using the task priority data in remote priority capture logic 32 to select one of processors P0, P1, P2, and P3 to receive the interrupt. A central agent 44 includes encode/decode logic 36, remote priority capture logic 32, and lowest priority logic 42. The redirected interrupt message is provided through conductors 48 to encode/decode logic 36.

Processors P0, P1, P2, and P3 have identification numbers, e.g., APIC IDs. The APIC ID may be, for example, supplied at power on or reset. Lowest

priority logic 42 provides the selected APIC ID number with the interrupt message. The interrupt message is provided through encode/decode logic 36 to processor bus 18 and the selected processor. Encode/decode logic of the selected processor recognizes the APIC ID number and passes the interrupt message. The interrupt message with associated bits (e.g., APIC ID number) may be passed in only one or in more than one phase or packet.

Where two or more processors have an equal lowest priority, lowest priority logic 42 may select the processor based on, for example, highest or lowest processor APIC ID, or in a round robin basis. The term "lowest priority" does not require that there be more than two different priority values. For example, if there is only one value of task priority data, it is the lowest.

Examples of the interrupt messages or other interrupt signals that may be provided over processor bus 18 include interrupt destination and vector signals, interrupt acknowledge signals, end-of-interrupt (EOI) signals, interprocessor interrupt (IPI) messages, other control signals or a combination of these signals. Some interrupt messages are not lowest priority signals and should not be redirected.

A lowest priority interrupt message may be provided to a processor destination by an interrupt controller, other circuitry, or the operating system (OS). In such a case, lowest priority logic 42 provides a destination redirection. However, the destination selected by lowest priority logic 42 may be the same as

the original destination, because the original destination happened to be the lowest priority processor. Accordingly, redirection does not mean a different direction, but a direction provided at a later stage. Alternatively, in the case of lowest priority interrupt messages, the processor destination may be provided for the first time by lowest priority logic 42.

The processor may provide the signals that are representative of task priority to processor bus 18 under the initiative of the processor providing the signal or under the request of external logic. For example, in a first embodiment, processors P0, P1, P2, and P3 provide signals representative of their task priority designations to processor bus 18 each time the task priority designation is changed. In a second embodiment, processors P0, P1, P2, and P3 provide signals representative of their task priority designations in response to a request from lowest priority logic 42 or an interrupt controller in response to their receiving a lowest priority interrupt message. In a third embodiment, remote priority capture logic 32 periodically requests an update of the processors. Other mechanisms may be used for updating the task priority data in remote priority capture logic 32. The operating system or other software may direct the updates.

Referring to FIG. 2, as an example, processor P0 includes a local APIC 52 which includes a local TPR (LTPR) 54. APIC 52 is an example of structure within interrupt control logic 22. In one embodiment, LTPR 54 holds an 8 bit task priority designation, the first 4 MSBs of which specify 16 priority classes.

In other embodiments, LTPR 54 could have a greater or lesser number of bits, or bits with different or additional significance. Referring to FIGS. 1, 2, and 3, encode/decode logic 58 includes encode logic that encodes, for example, the 4 MSBs of LTPR 54 onto a signal for processor bus 18 to be decoded by encode/decode logic 36, which provides signals on conductors 38 to remote priority capture logic 32.

Referring to FIGS. 2 and 3, one embodiment of remote priority capture logic 32 includes remote task priority registers (RTPRs) 62, 64, 66, and 68. RTPR 62 holds task priority data indicative of a task priority specified in LTPR 54 of processor P0. RTPRs 64, 66, and 68 hold task priority data indicative of task priorities specified in LTPRs (not shown) of processor P1, P2, and P3 respectively.

Referring to FIG. 4, as an example, RTPR 62 includes four bits (e.g., bits 0 - 3) that holds task priority data that is indicative of a task priority specified in LTPR 54 of processor P0, if processor P0 is available for LPIDA. These four bits in RTPR 62 do not have to be identical to the four MSBs of the LTPR. For example, they could be inverted. RTPR 62 also includes a bit (e.g., bit 7), which indicates whether processor P0 is available for LPIDA. In the particular embodiment, RTPRs 64, 66, and 68 each also include four bits that hold task priority data that is indicative of task priorities specified in LTPRs (not shown) in processors P1, P2, and P3, respectively, if processors P1, P2, and P3 are

available for LPIDA. RTPRs 64, 66, and 68 also include an enable/disable bit that indicates whether processors P1, P2, and P3 are available for LPIDA. The enable/disable bit is enabled with a first voltage level (e.g., a logical high voltage), indicating a processor is available for LPIDA. The enable/disable bit is disabled with a second voltage level (e.g., a logical low voltage), indicating a processor is not available for LPIDA.

The task priority data in the RTPR is indicative of the task priority designation of the LTPR even though the RTPR task priority data is not identical to the task priority designation in the LTPR. For example, in an embodiment described above, the RTPR holds the 4 MSBs of the 8 bit task priority designation of a corresponding LTPR. However, for purposes of this invention, the 4 MSBs of an 8-bit number are considered to be indicative of the entire 8 bit number. The 4 MSBs are comprehensive enough to achieve lowest priority semantics. In other words, the 4 LSBs of an 8-bit task priority designation are not significant enough to matter for purposes of this invention. Of course, the task priority data in the RTPR could include all the bits of the LTPR.

Further, depending on the implementation, there is some chance that the task priority data in task remote priority capture logic 32 will sometimes not exactly reflect the actual task priority designations of the processors available for LPIDA, because the task priority designations change with time. However, the task priority data is still indicative of the task priorities of the processors

available for LPIDA, even if the indication is not always perfect or LPIDA does not always select the processor with the lowest priority. In the case in which a processor is not available for LPIDA because the enable/disable bit is set to disable, in one embodiment, the task priority data in that RTPR is updated just as if the enable/disable bit were set to enable. In another embodiment, the task priority data is not updated until the enable/disable bit will be set to enable. In that last mentioned embodiment, the task priority data might not be indicative of a task priority of the corresponding processor, although it does not matter because it is not used in LPIDA. In that embodiment, the contents of the four bits is not updated. Of course, in the case in which a processor is not available for LPIDA because the processor is not active or missing from the system, the contents of the four task priority data bits of the corresponding RTPR will be meaningless.

If one of the processors of a multi-processor system is not present in the system, at a given APIC ID, the enable/disable bit is disabled in the corresponding RTPR. Under one embodiment, enable/disable bit of the RTPR needs to be set the first time the RTPR is accessed (updated by the appropriate processor) and, once set, must remain set until a 'cold' reset event occurs. The RTPR can be updated based on a number of event types. Two of the possible options are: (1) direct BIOS access to the RTPR or (2) a RTPR update special cycle transaction (an example of which described in connection with FIG. 7) by the corresponding

agent, controls the state of the enable/disable bit. An upgrade/downgrade of the RTPR may occur as a result of Power-On Self Test (POST) before an I/O interrupt enters the system. The processor may also raise its priority to the highest level to avoid an interrupt.

In the illustrated example of FIG. 4, RTPR 62 includes additional bits (e.g., bits 4-6). In one embodiment of the invention, the additional bits are not used and are reserved. In another embodiment of the invention, one or more of the bits may be used for various purposes. In still another embodiment, there are no additional bits in the RTPRs. The enable/disable function may be accomplished with two bits rather than one.

Referring to FIG. 3, as an example, lowest priority logic 42 may include a buffer 74 and analyzing logic 76. Analyzing logic 76 has access to the contents of the RTPRs through conductors 72. Analyzing logic 76 performs LPIDA to determine which of the participating RTPRs has the lowest priority (which may include resolving any ties in lowest priority). An optional buffer 74 may hold the lowest priority interrupt message until LPIDA is completed. A signal on conductor 86 indicates the APIC ID number or other indication of the selected processor, which is provided through encode/decode logic 36 to bus 18. The APIC ID number or other indication may be provided in a variety of forms to bus 18, and may be in the same or a different phase or packet than other information of the interrupt message.

Lowest priority logic 42 may use any of various known techniques to determine which of the participating RTPRs has the lowest value (or highest value if a logical 1 value is a lower priority that a logical 0). For example, lowest priority logic 42 could eliminate RTPRs having a logic 1 value in the MSB, and then eliminate RTPRs having bits having a logic 1 value in the second MSB and so forth. Lowest priority logic 42 could subtract values to see which is greater based on whether the result is positive or negative, or use various other methods. In most, if not all of the techniques, lowest priority logic 42 will select the processor much faster than in the case of APIC serial bus arbitration.

There may be circuitry (not shown in FIG. 1) between encode/decode logic 36 and remote priority capture logic 32, and between encode/decode logic 36 and lowest priority logic 42. Remote priority capture logic 32 and lowest priority logic 42 are not required to be in the processor bus bridge (which in some embodiments is called the North bridge). FIG. 5 illustrates a system 100 wherein central agent 44 is included in a processor bridge (or chipset) 104. Bridge 104 interfaces between an I/O bus 108 and peripherals 112A and 112B connected thereto (which may interface according to a well known Peripheral Component Interconnect (PCI) standard). Peripherals 112A and 112B represent a variety of components including interrupt controllers or bridges to other busses. Bridge 104 may be designed so that the features of the present invention are transparent to peripherals and/or operating system software. That is, under one embodiment,

the peripheral and/or operating system need not know whether a processor bus or APIC serial bus are being used to communicate between the processors and bridge.

FIG. 5 illustrates one of the variety of ways of implementing bridge 104. An I/O interrupt controller 114 may be constructed according to a well known manner or be especially designed for the present invention. Interrupt controller 114 may include an I/O redirection table to provide a relationship between I/O interrupt requests and the destination of the targeted request. I/O redirection table may provide interrupt vectors to identify the entry into a table which designates the appropriate interrupt service routine. An inbound queue 120 holds interrupt requests waiting to be sent to a processor. An optional outbound queue 126 holds signals communicated from a processor.

System 100 of FIG. 5 does not include an APIC serial bus. Referring to FIG. 6, a system 170 includes a bridge 174 that includes remote priority capture logic and redirection logic according to an embodiment of the present invention. Bridge 174 allows interrupt messages to pass on processor bus 18 between bridge 174 and processors P0, P1, P2, and P3. System 170 also includes an APIC serial bus 178 that allows operations that are performed by APIC serial busses of the prior art. Bridge 174 may therefore be used by processors that understand interrupt messages on processor bus 18 and by processors that understand interrupt messages on APIC serial bus 178. Depending on the processor, it may

be that the direct interface with the processor will have to be different, but a common bridge could be used by both.

The various bridges illustrated and discussed may include a variety of components that are well known in the art, but which are not illustrated and discussed here because such illustration and discussion are not necessary to understand the present invention.

Lowest priority logic 42 may be used in connection with IPI messages. For example, the IPI signal from the directing processor P0, P1, P2, or P3 is forwarded to bridge 104 or 170. Merely as an example, the IPI message may be forward to I/O bus 108 and then directed through bridge 104 back to the processor selected by lowest priority logic 42. Alternatively, the IPI signal may be forwarded directly to the inbound queue 120. When the IPI is first provided to bus 18, an address bit (e.g., Aa3#) in the first phase of the IPI may be set to a first voltage (e.g., high) indicating that interrupt request is to be ignored by the processors, but is to be consumed by the bridge. When the IPI request returns from bridge, the bit will be set to a second voltage (e.g., low), so that the selected (target) processor will consume the IPI.

The following table summarizes the effect of the state of certain signals on processor bus 18 in one embodiment of the invention, where X means don't care; Ab5# and Ab6# are in the second phase of the transactions; during fixed delivery

mode, lowest priority logic 42 does not perform LPIDA; and during redirected delivery mode, lowest priority logic 42 does perform LPIDA:

Aa3#	Ab6# (EXF3#)	Ab5# (EXF2#)	Interrupt Transaction Type
0	0	0	Fixed Delivery Mode - Physical Destination Mode
0	0	1	Fixed Delivery Mode - Logical Destination Mode
0	1	X	Reserved
1	0	0	Redirected Delivery Mode - Physical Destination Mode
1	0	1	Redirected Delivery Mode - Logical Destination Mode
1	1	0	Reserved
1	1	1	End of Interrupt (EOI)

A bit (e.g., Aa3#) in the address field of the interrupt message may indicate whether LPIDA should occur at all for a particular interrupt message, regardless of the states of the enable/disable bits in remote priority capture logic 32. This bit, which may be called the redirection bit, may be computed based on the lowest priority encoding, bits [10:8] (e.g., 001), of an I/O redirection table in interrupt controller 114.

Lowest priority logic 42 may be used in connection with physical destination mode and logical destination mode. Under one embodiment, under physical destination mode, target processors are selected based on unique APIC IDs. Accordingly, each interrupt can be directed to a given processor based on its unique APIC ID value. Under logic destination mode, target processors are

selected based on a logical ID value programmed into each APIC. Since a logical ID is programmed, and therefore not necessarily unique to a given processor, they can identify a group of processors to be targeted. Interrupt message bits (e.g., Ab5# and Ab6#) may indicate whether a physical or logical destination mode is used. In physical destination mode, lowest priority logic 42 may select any of the processors on the cluster as the processor to receive the interrupt (assuming the enable/disable bit in remote priority capture logic 32 is set to enable for that processor).

In logic destination mode, the system may operate as follows. Lowest priority logic 42 or other circuitry checks the logical ID to determine whether the interrupt is directed to a processor within the logical cluster. If the interrupt message is directed to one of the processors on processor bus 18, lowest priority logic 42 may determine the destination processor from the group of processors indicated in the logical ID. The directed interrupt on processor bus 18 will be sent to the processor with the lowest interrupt priority from the group of processors indicated by the logical ID. For example, suppose four processors are in a system at the logical mode cluster address of 00xxh (hex). If an I/O interrupt arrives at the host bridge and has the logical ID of '00000111', and is tagged to be redirected, LPIDA may be determined for P2 to P0.

FIGS. 1 illustrates multi-processor systems. Alternatively, central agent 44 or bridge 100 may be used in connection with a single processor. In that case,

in one embodiment, lowest priority logic 42 always sends interrupt messages to that processor. Under one approach, remote priority capture logic 32 is inactive if there is only one processor. Under another approach, remote priority capture logic 32 is active, but the RTPR corresponding to the processor is the only enabled RTPR. The only processor in the system may or may not provide signals representative of its task priority. Under one embodiment, if only one enable/disable bit is set in remote priority capture logic 32, lowest priority logic 42 will direct the interrupt to that processor regardless of what is in the TPR field. Alternatively, remote priority capture logic 32 could include some other indication as to there being only one processor. Where there is only one processor, the priority captured by remote priority capture logic 32 may merely be that a processor is available for interrupts.

According to one embodiment of the invention, interrupt messages are assigned a memory address within a one megabyte space of memory. In a 4 Gigabyte space, the one megabyte location can be between FEE00000h and FEEFFFFFh. The memory location may be used to identify a particular destination.

In one embodiment, processors P0, P1, P2, and P3, and encode/decode logic 36, and (optionally) the operating system are designed so that processors P0, P1, P2, or P3 can write RTPR updates directly to RTPRs 62, 64, 66, or 68, respectively. In this embodiment, the RTPRs may be treated as I/O space. In

another embodiment, processors P0, P1, P2, and P3, encode/decode logic 36, and (optionally) the operating system do not allow that capability for processors P0, P1, P2, and P3 to write a RTPR update directly to RTPRs 62, 64, 66, and 68, but rather use a RTPR update special cycle transaction over processor bus 18 to update the RTPRs. Processors of this other embodiment are particularly suited for currently used operating systems and interrupt semantics.

Referring to FIG. 7, one embodiment of a RTPR update special cycle transaction includes two phases 182 and 184. Phase 182 includes a command field (e.g., 5 LSBs) and an address field (e.g., 26 MSBs). As an example, bits "01000" in the command field indicate a special cycle. In the case of a special cycle, the address bits may be don't care. Phase 184 includes a byte enable field (e.g., 00001000 or 08h) that indicates an RTPR update cycle; a processor ID field that indicates which processor is providing the update; an enable/disable (E/D) bit to indicate whether the processor is available for LPIDA; and TPR bits representing, for example, the four MSBs of the corresponding LTPR. In the example of FIG. 4, the TPR bits may be placed in bits 0-3 and the E/D bit may be placed in bit 7 of the RTPR. The E/D bit and TPR bits may be provided within what is otherwise the 8-bit attribute field.

Encode/decode logic 36 responds to the command field of phase 182 and the byte enable field of phase 184 by providing an update to the RTPR designated by the processor ID field in remote priority capture logic 32. The

RTPR is updated with bits representing the E/D bit and/or the TPR bits. Where the E/D bit indicates the processor is disabled, the RTPR may or may not be also updated with the TRP bits. In one embodiment, where the E/D bit indicates the processor is disabled, the processor does not provide meaningful task priority data in the TPR bits. In another embodiment, the processor provides current TPR bits regardless of the state of the E/D bit.

Through the signals of FIG. 7, processor P0, P1, P2, or P3 and encode/decode logic 36 provide a hardware assist mechanism to alias RTPR 62, 64, 66, or 68 without the operating system being aware of the update. (Alternatively, the operating system could be aware.) Various other signal arrangements may be used in place of those illustrated in FIG. 7. For example, all the information could be provided in one phase. As another example, phase 184 could provide an update for more than one of RTPR 62, 64, 66, and 68 at a time. In the illustrated and described example, the RTPRs hold only four bits to represent the processor task priority. The TPR bits in phase 184 may represent more or less than the four MSBs of the corresponding LTPR, where the RTPRs holds more or less than four bits, respectively, to represent the task priority.

Additional Information and Embodiments

The specification does not describe or illustrate various well known components, features, and conductors, a discussion of which is not necessary to understand the invention and inclusion of which would tend to obscure the

invention. Furthermore, in constructing an embodiment of the invention, there would be various design tradeoffs and choices, which would vary from embodiment to embodiment. Indeed, there are a variety of ways of implementing the illustrated and unillustrated components.

The borders of the boxes in the figures are for illustrative purposes and do not restrict the boundaries of the components, which may overlap. The relative size of the illustrative components does not suggest actual relative sizes. Arrows show principle data flow in one embodiment, but not every signal, such as requests for data flow. As used herein "logic" does not mean that software control cannot be involved. The term "conductor" is intended to be interpreted broadly and includes devices that conduct although they also have some insulating properties. There may be intermediate components or conductors between the illustrated components and conductors.

The phrase "in one embodiment" means that the particular feature, structure, or characteristic following the phrase is included in at least one embodiment of the invention, and may be included in more than one embodiment of the invention. Also, the appearances of the phrase "in one embodiment" in various places in the specification are not necessarily all referring to the same one embodiment.

Lowest priority logic 42 may direct (or redirect) interrupts across multiple nodes.

A processor could have more than one priority for different kinds of tasks, and the remote priority capture logic and lowest priority logic could take the different priorities into account.

The encode and decode logic of encode/decode logic 36 may be physically connected or separated. The encode and decode logic of encode/decode logic 58 may be physically connected or separated.

For a multiprocessor system within a single chip, there could be interrupt capture logic and lowest priority logic within that chip.

The term "connected" and "coupled" and related terms are used in an operational sense and are not necessarily limited to a direct connection or coupling. If the specification states a component or feature "may", "can", "could", or is "preferred" to be included or have a characteristic, that particular component or feature is not required to be included or have the characteristic. The term "responsive" includes completely or partially responsive.

Those skilled in the art having the benefit of this disclosure will appreciate that many other variations from the foregoing description and drawings may be made within the scope of the present invention. Accordingly, it is the following claims including any amendments thereto that define the scope of the invention.

CLAIMS

What is claimed is:

1. An apparatus for use in connection with computer system, the apparatus comprising:

remote priority capture logic to hold task priority data indicative of a task priority of each processor in the computer system that is available for lowest priority interrupt destination arbitration (LPIDA); and

lowest priority logic to perform the LPIDA to select which processor in the computer system is to receive an interrupt message based on contents of the remote priority capture logic.

2. The apparatus of claim 1, wherein there this is only one processor in the computer system.

3. The apparatus of claim 1, wherein there are multiple processors in the system and the availability of the processors for LPIDA varies with time.

4. The apparatus of claim 1, wherein the remote priority capture logic holds task priority data indicative of a task priority of each processor in the computer system that is not available, but the lowest priority logic only considers the indicative data corresponding to those processors that are available for LPIDA.

5. The apparatus of claim 1, wherein the remote priority capture logic includes an enable/disable bit to indicate whether a processor is available for the LPIDA.

6. The apparatus of claim 1, wherein the remote priority capture logic includes an enable/disable bit.

7. The apparatus of claim 6, wherein the remote priority capture logic is capable of holding task priority data, but does not actually hold task priority data where the computer system has only one processor, and wherein the lowest priority logic is capable of performing lowest priority data and selects the processor based on the state of the enable/disable bit.

8. The apparatus of claim 1, wherein the apparatus is a bridge.

9. The apparatus of claim 1, wherein the lowest priority logic uses the LPIDA to perform interrupt redirection.

10. The apparatus of claim 1, further including an interrupt controller that provides the interrupt message.

11. An apparatus for use in connection with a processor, the apparatus comprising:

remote priority capture logic to hold task priority data indicative of a task priority of the processor while the processor is available for lowest priority interrupt destination arbitration (LPIDA); and

lowest priority logic to perform the LPIDA to select the processor to receive an interrupt message based on contents of the remote priority capture logic.

12. The apparatus of claim 11, wherein the processor is the only processor associated with the lowest priority logic and the lowest priority logic always selects the processor to receive the interrupt message.

13. The apparatus of claim 11, wherein the remote priority capture logic includes an enable/disable bit to indicate whether a processor is available for the LPIDA.

14. An apparatus for use in connection a processor, the apparatus comprising:

remote priority capture logic to hold an indication that a processor is available to receive an interrupt; and

lowest priority logic to select the processor is to receive an interrupt message based on contents of the remote priority capture logic.

15. The apparatus of claim 14, wherein the indication is the setting of an enable/disable bit corresponding to the processor.

16. An apparatus for use in connection with processors of a multi-processor system, the apparatus comprising:

remote priority capture logic to hold task priority data indicative of task priorities of those of the processors that are available for lowest priority interrupt destination arbitration (LPIDA); and

lowest priority logic to perform the LPIDA to select which of the processors is to receive an interrupt message based on contents of the remote priority capture logic.

17. A multi-processor system, comprising:

processors;

a processor bus coupled to the processors;

remote priority capture logic to hold task priority data indicative of a task priority of the processors while they are available for lowest priority interrupt destination arbitration (LPIDA); and

lowest priority logic to perform the LPIDA to select which of the processors is to receive an interrupt message based on contents of the remote priority capture logic, the interrupt message being provided to the processor through the processor bus.

18. The system of claim 17, further including an interrupt controller that provides the interrupt message and wherein the lowest priority logic provides interrupt destination redirection through the LPIDA.

19. The system of claim 17, wherein at least one of the processors may be available for LPIDA during a time that at least another one of the processors is not available for LPIDA.

20. The system of claim 17, wherein the remote priority capture logic includes remote task priority registers (RTPRs) to hold the task priority data.

21. The system of claim 17, wherein the number of RTPRs exceeds the number of the processors.

22. A multi-processor system, comprising:

processors;

a processor bus coupled to the processors;

encode/decode logic coupled to the processor bus;

remote priority capture logic to hold task priority data indicative of a task priority of the processors while they are available for lowest priority interrupt destination arbitration (LPIDA); and

lowest priority logic to perform the LPIDA to select which of the processors available for LPIDA is to receive an interrupt message based on contents of the remote priority capture logic, the interrupt message being provided through the encode/decode logic to the selected processor through the processor bus.

23. The system of claim 22, wherein the lowest priority logic provides a signal to the interrupt message to indicate the selected processor.

24. The system of claim 22, further including an interrupt controller that provides the interrupt message and wherein the lowest priority logic provides interrupt destination redirection through the LPIDA.

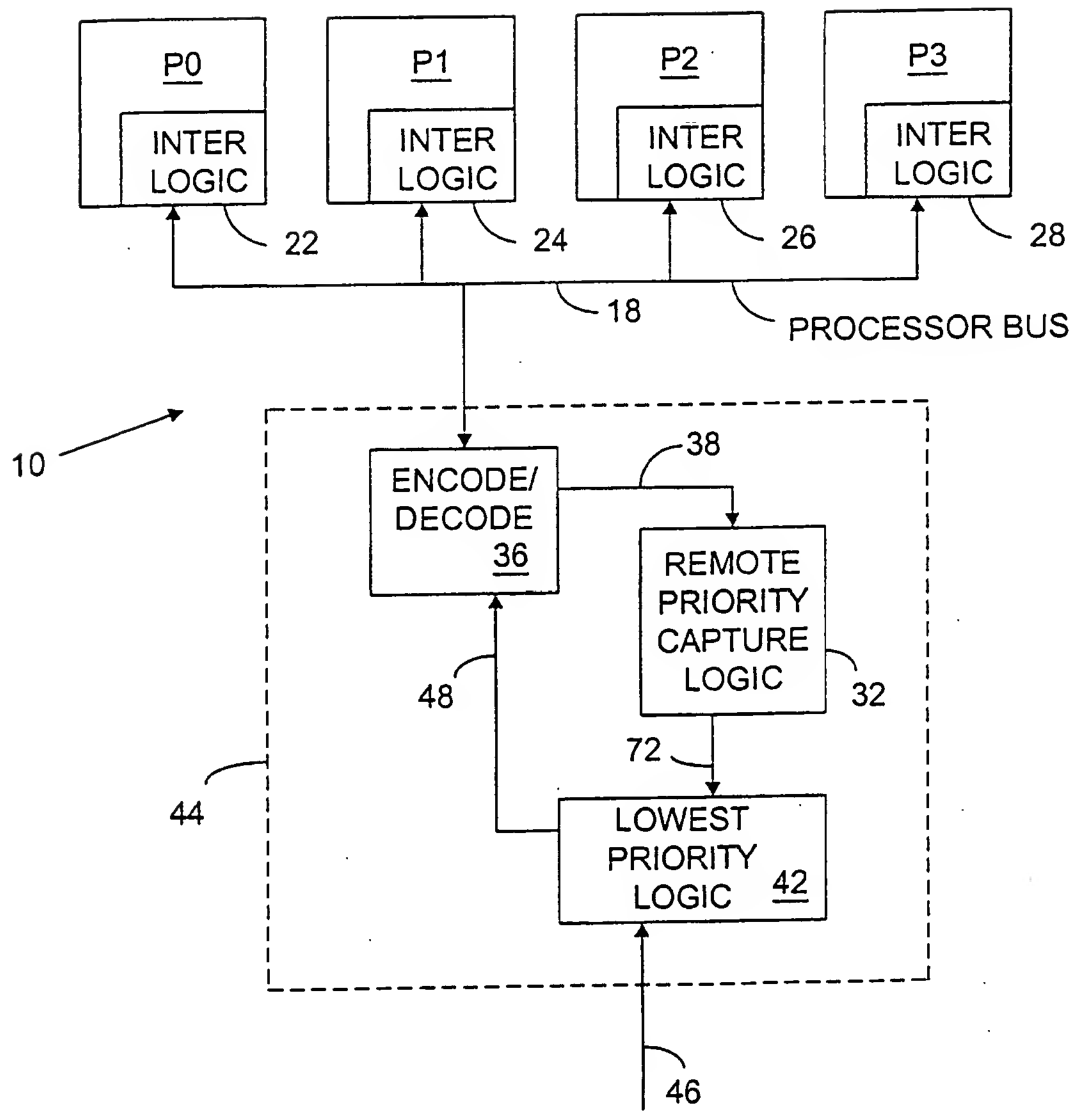

P0
INTER LOGIC
22
P1
INTER LOGIC
24
P2
INTER LOGIC
26
P3
INTER LOGIC
28
18
PROCESSOR BUS
10
ENCODE/ DECODE 36
38
REMOTE PRIORITY CAPTURE LOGIC
32
48
72
44
LOWEST PRIORITY LOGIC 42
46

FIG. 1

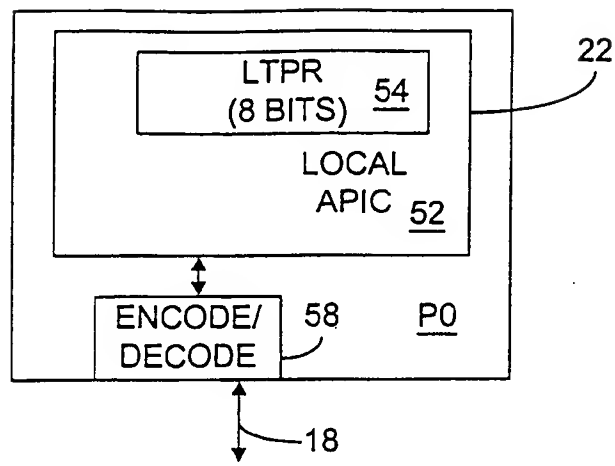

LTPR
(8 BITS) 54
LOCAL
APIC 52
22
ENCODE/
DECODE
58
P0
18

FIG. 2

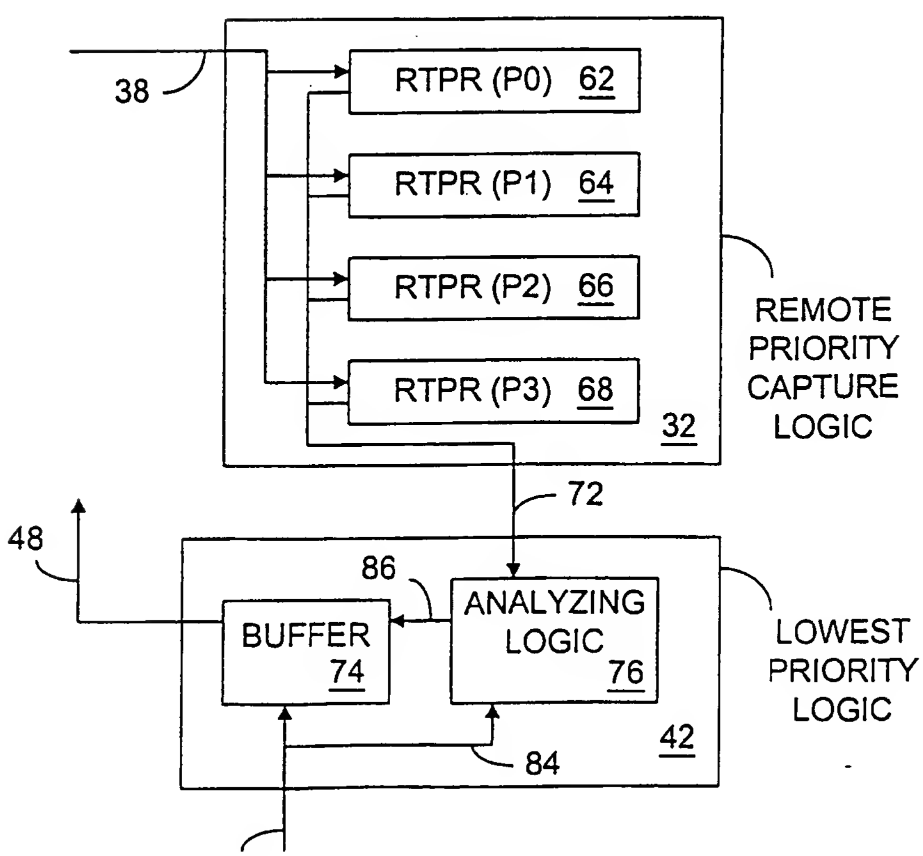

38
RTPR (P0) 62
RTPR (P1) 64
RTPR (P2) 66
RTPR (P3) 68
32
REMOTE
PRIORITY
CAPTURE
LOGIC
72
48
86
BUFFER
74
ANALYZING
LOGIC
76
42
84
LOWEST
PRIORITY
LOGIC
46

FIG. 3

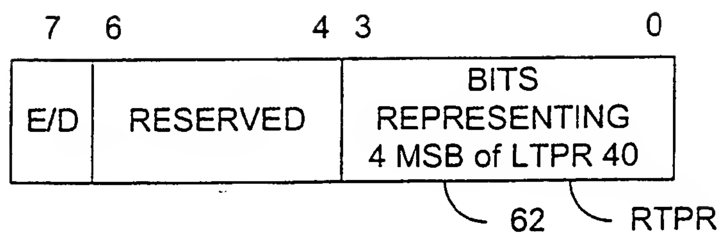
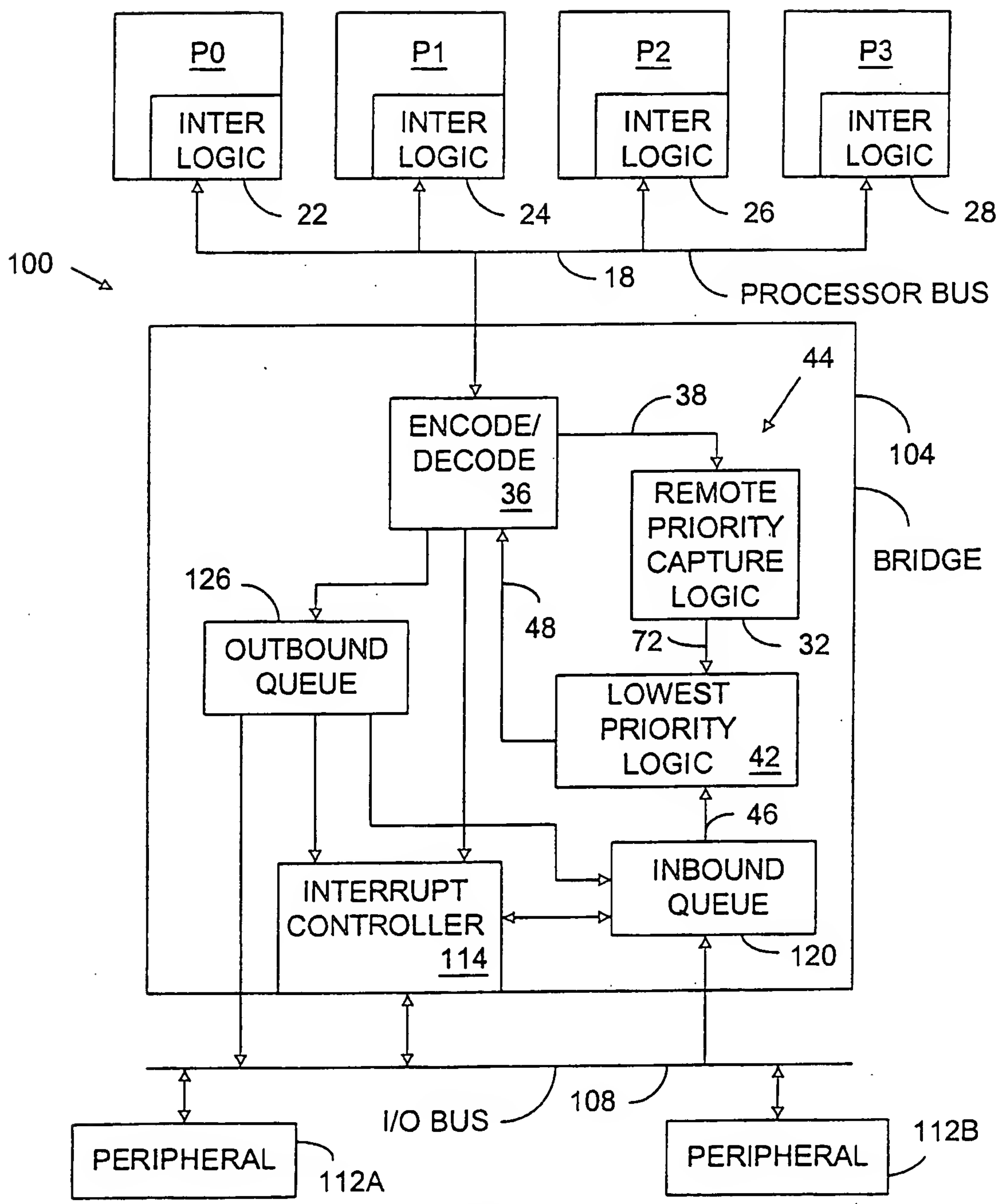


FIG. 4



P0
INTER LOGIC
22
P1
INTER LOGIC
24
P2
INTER LOGIC
26
P3
INTER LOGIC
28
100
18
PROCESSOR BUS
44
38
ENCODE/ DECODE 36
104
REMOTE PRIORITY CAPTURE LOGIC
BRIDGE
126
48
OUTBOUND QUEUE
72
32
LOWEST PRIORITY LOGIC 42
46
INTERRUPT CONTROLLER 114
INBOUND QUEUE
120
I/O BUS
108
PERIPHERAL
112A
PERIPHERAL
112B

FIG. 5

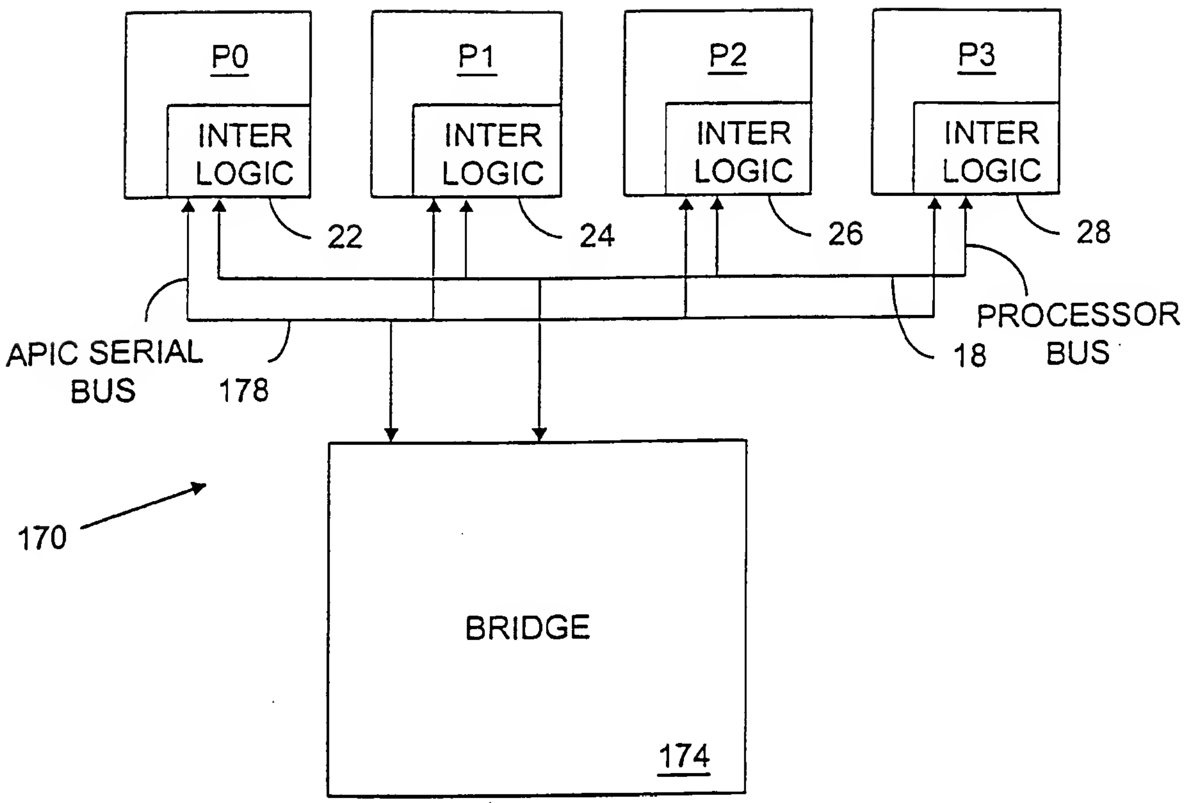

P0
INTER LOGIC
P1
INTER LOGIC
P2
INTER LOGIC
P3
INTER LOGIC
22
24
26
28
APIC SERIAL BUS
178
18
PROCESSOR BUS
170
BRIDGE
174

FIG. 6

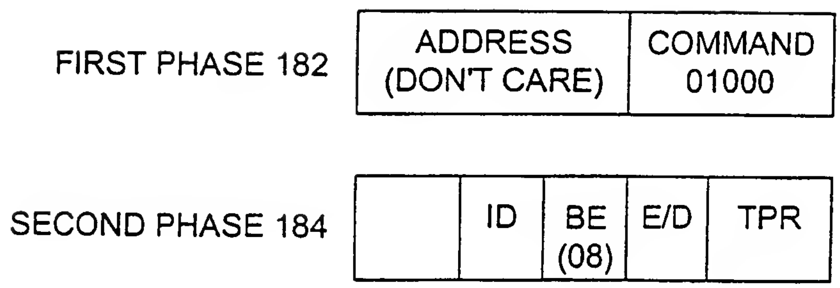

FIRST PHASE 182
ADDRESS (DON'T CARE)
COMMAND 01000
SECOND PHASE 184
ID
BE (08)
E/D
TPR

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/23354

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :G06F 13/26
US CL :710/263
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 710/263, 261, 262, 264, 267

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,696,976 A (NIZAR et al.) 09 December 1997, figure 2, 5; col. 13, lines 23-30; col. 15, lines 6-38; col. 4, lines 53-59; col. 15, lines 32-37; col. 13, line 38 through col. 14, line 15; col. 8, lines 42-47.	1-24

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search	Date of mailing of the international search report
05 MARCH 1999	14 MAY 1999

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	AYAZ SHEIKH James R. Matthews
Facsimile No. (703) 305-3230	Telephone No. 703-305-9648